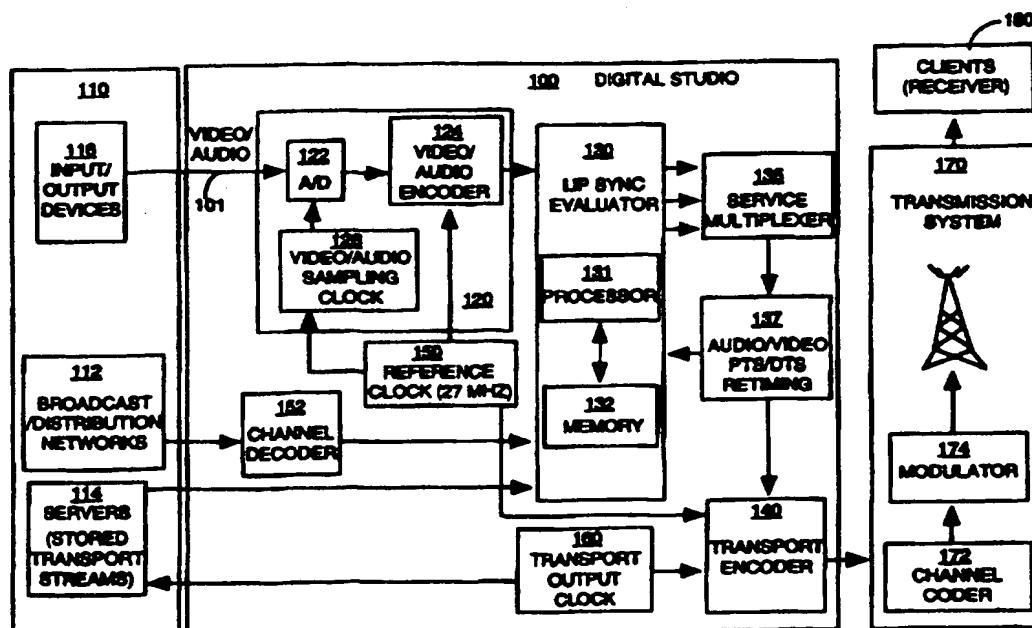


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>H04N 9/475</b>		A1	(11) International Publication Number: <b>WO 97/46027</b>
			(43) International Publication Date: 4 December 1997 (04.12.97)
(21) International Application Number: PCT/US97/10186 (22) International Filing Date: 28 May 1997 (28.05.97) (30) Priority Data: 60/018,554                      29 May 1996 (29.05.96)                      US (71) Applicant: SARNOFF CORPORATION [US/US]; 201 Washington Road, CN-5300, Princeton, NJ 08543-5300 (US). (72) Inventors: FOX, Edward, C.; 4 Steele Drive, Cranbury, NJ 08512 (US). LYONS, Paul, W.; 9 Kenyon Drive, New Egypt, NJ 08533 (US). WINE, Charles, M.; 144 Hamilton Avenue, Princeton, NJ 08540-3858 (US). (74) Agent: BURKE, William, J.; Sarnoff Corporation, 201 Washington Road, CN-5300, Princeton, NJ 08543-5300 (US).			(81) Designated States: CA, JP, KR, MX, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.          Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

BEST AVAILABLE COPY

(54) Title: PRESERVING SYNCHRONIZATION OF AUDIO AND VIDEO PRESENTATION



## (57) Abstract

A method and apparatus (100) for preserving audio and video presentation synchronization, i.e., lip sync, during the splicing operation by selectively deleting, if necessary, an audio/video access unit to avoid overlapping of audio/video frames in the spliced output stream.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## PRESERVING SYNCHRONIZATION OF AUDIO AND VIDEO PRESENTATION

5 This application claims the benefit of U.S. Provisional Application No. 60/018554 filed May 29, 1996.

This invention was made with U.S. government support under contract number 70NANB5H1174. The U.S. government has certain rights in this invention.

10 The present invention relates to an apparatus and concomitant method for preserving audio and video presentation synchronization when splicing data streams, e.g., transport streams, from one or more sources. More particularly, this invention relates to a method and apparatus that preserves audio and video presentation synchronization during the  
15 splicing operation by selectively deleting, if necessary, an audio/video access unit to avoid overlapping of audio/video frames in the spliced output stream.

### BACKGROUND OF THE INVENTION

20 Typically, each data stream, when in transport format, carries a plurality of audio and video data streams (substreams), e.g., MPEG system layers define Packetized Elementary Streams (PES) which may carry encoded audio and video streams. Furthermore, MPEG provides a mechanism for time stamping the individual elementary stream  
25 components of a program with Presentation Time Stamps (PTS) in the PES layer for time synchronization between the video and audio components (program components) at the time of origination.

However, the presentation time of the various program components are not synchronous to each other but are synchronized to the system  
30 clock, e.g., a 27 MHz reference clock. Specifically, the audio and video presentation units have different durations. An audio presentation unit or frame is fixed at 32 msec, while the video presentation unit or frame varies with video format and is not fixed at 32 msec. Maintaining synchronization between the video signal and the associated audio signal  
35 is vital in providing high quality presentations, i.e., "lip sync". Lip sync is

the synchronization of audio and video presentation, e.g., the synchronization of a soundtrack consisting of dialogue, music, and effects with the pictures of a program.

This requirement creates a problem when switching from one program to another program during a splicing or switching operation. The video and audio units are typically not aligned in the time domain. Thus, switching encoded data streams, e.g., at either a video or an audio "access unit" (a coded representation of a video or an audio presentation unit) creates a partial access unit in the other associated elementary stream that was not aligned at the switch point, e.g., aligning the video access units of two data streams may cause overlap of their audio access units and vice versa.

However, if one attempts to align both the video and the audio by creating a continuous flow of access units for both video and audio, the audio to video time relationships are disturbed causing them to lose synchronization.

Therefore, a need exists in the art for a method and apparatus for preserving audio/video lip sync when splicing data streams from multiple sources.

20

### SUMMARY OF THE INVENTION

The present invention is a method and apparatus for preserving audio and video presentation synchronization, i.e., lip sync, when splicing data streams from one or more sources. The invention preserves audio/video lip sync during the splicing operation by selectively deleting, if necessary, an audio/video access unit to avoid overlapping of audio/video frames in the spliced output stream.

25

### BRIEF DESCRIPTION OF THE DRAWINGS

The teachings of the present invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

FIG. 1 illustrates a block diagram of a digital studio system employing the present invention;

FIG. 2 illustrates a splicing operation that is premised on maintaining video presentation continuity;

FIG. 3 illustrates a splicing operation that is premised on maintaining video presentation continuity, where an overlap condition  
5 does not exist;

FIG. 4 illustrates an alternative embodiment of the present invention where the splicing operation is premised on maintaining audio presentation continuity; and

FIG. 5 illustrates a flowchart of a method for preserving audio/video  
10 lip sync when splicing a plurality of data streams into an output stream.

To facilitate understanding, identical reference numerals have been used, where possible, to designate identical elements that are common to the figures.

15

### DETAILED DESCRIPTION

FIG. 1 illustrates a block diagram of a communication environment having a digital studio 100, a source section 110, a transmission system 170 and a plurality of clients 180. Alternatively, those skilled in the art will realize that the digital studio may comprise the source section 110 (or  
20 portion thereof) and the transmission system 170.

Generally, the digital studio serves to distribute various programs to a plurality of clients/receivers 180. In interactive mode, the digital studio 100 permits the clients 180 to selectively request and/or control the various studio resources within source section 110. Each of the clients may  
25 include, but is not limited to, a set-top terminal, a receiver, a computer or a storage device. In fact, since it is contemplated that the digital studio will be under distributed control, other studio components may serve as clients as well, e.g., the various program sources within source section 110.

30 Source section 110 comprises a plurality of program sources, e.g., broadcast/distributing networks or devices 112, servers 114 and various input/output devices 116. More specifically, the broadcast/distribution devices or networks 112 may include, but are not limited to, a satellite distribution network, a broadcast network, a local "live-feed" network or  
35 even another digital studio. These devices may generate a transport

stream that contains full-motion video, i.e., a sport event with a large quantity of motion and detail.

Although the present invention is described below with reference to transport streams, it should be understood that the present invention can be applied to other bitstream formats, including but not limited to, MPEG program streams or bitstreams in accordance with the asynchronous transfer mode (ATM). Furthermore, although the present invention is described below with reference to a digital studio, it should be understood that the present invention can be adapted to other devices, e.g., playback devices such as a receiver or a video player.

Similarly, the servers 114 may include, but are not limited to, file servers holding a plurality of film and video sources, e.g., a movie (24 frames/second), a video (30 frames/second) of a lecturer or a video of a commercial. In turn, the input/output devices 116 may include, but are not limited to, monitors, various filters, transcoders, converters, codecs, cameras, recorders, interface devices and switchers. Each of the various studio components may incorporate the necessary hardware (e.g., one or more processors, computers or workstation) to store or implement software routines or objects.

In brief, these various program sources generate and/or store transport streams (or data streams in general which are processed by the studio) that are received and multiplexed (splicing operation) by a service multiplexer (splicer) 135 into a single bitstream, e.g., an output stream. This output stream is then encoded and packetized by a transport encoder 140, if not previously in transport format, to produce an output transport stream. A detail discussion of the splicing operation and its effect on audio/video lip sync are further disclosed below with reference to the digital studio 100 and FIGs. 2-4.

In turn, the output transport stream is forwarded to a channel coder 172, where error correction coding is applied. The modulator 74 then modulates the error-coded output transport stream onto a carrier signal, using one of many possible modulation schemes, e.g., 8-vestigial sideband modulation (VSB), 16-VSB, Quadrature Amplitude Modulation (QAM) and the like.

As discussed above, the significant differences in the timing information embedded in the various transport streams pose a significant challenge in preserving audio/video lip sync when splicing transport streams from multiple sources. FIG. 1 illustrates a block diagram of a digital studio 100 that addresses this timing criticality by incorporating a lip sync evaluator 130.

More specifically, FIG. 1 illustrates the digital studio 100 incorporating a lip sync evaluator 130 that receives input transport streams from various program sources 112-116. Since the program sources may generate data streams with different formats, the digital studio 100 may employ the necessary devices to reformat or decode the various data streams into a common format prior to performing the splicing operation.

To illustrate, an optional encoding section 120 is employed to generate or convert various input signals into MPEG compliant elementary streams. For example, one or more of the input devices 116 may forward "raw" input signals such as analog video and audio signals from a camera to the studio 100 via path 101. The input signals are sampled and converted by an analog-to-digital (A/D) converter 122 into a digitized signal. The sampling frequency for the A/D is provided by the video/audio sampling clock 126, which, in turn, is frequency locked to the reference clock 150. The digitized signal is then encoded by video/audio encoder 124, to implement various video and audio compression methods, i.e., source encoding. The resulting transport stream is then ready to be forwarded to the lip sync evaluator 130 for evaluation before the transport stream is spliced into an output stream.

Similarly, an optional channel decoder 152 is employed to decode or convert various input signals into MPEG compliant transport streams. Since program sources may include broadcast/distribution networks 112, the data streams from such networks are typically channel encoded. Thus, the channel decoder 152 removes the channel coding prior to forwarding the transport streams to the lip sync evaluator.

Finally, the digital studio 100 may receive data streams that are already in the proper stream format from a server 114. In such case, the transport streams are read from the server using transport output clock

160 and are forwarded directly to the lip sync evaluator for evaluation.

Although the preferred embodiment of the present invention performs the splicing operation at the transport stream layer, the present invention can be adapted to splicing operations performed at other "stream layers", e.g., at the elementary stream layer. However, performing the splicing operation at lower stream layers may require more computational overhead, thereby incurring additional delay in the splicing operation.

Returning to FIG. 1, the input transport streams to be spliced are passed through a lip sync evaluator 130 that preserves the audio/video lip sync in the output transport stream. More specifically, the lip sync evaluator 130 determines the spacing between audio or video access units at the splice point. If the spacing indicates an overlap situation, then one of the overlapping audio or video access unit is deleted or dropped from the spliced output stream. The lip sync evaluator can be implemented using a processor 131 with an associated memory 132 or as a software implementation residing in a memory operated by a studio controller (not shown).

Furthermore, a portion of the memory 132 can be used as a buffer to temporally store relevant access units to determine whether an overlap condition exists. If an overlap condition is detected, then one of the overlapping access unit will be deleted from the memory without being inserted into the output transport stream as discussed below.

Referring to FIGs. 2-4, each of these figures illustrates the timing relationship between the video access units and their associated audio access units for three different transport streams, i.e., a first input transport stream representative of a first program, a second input transport stream representative of a second program and the resulting output (spliced) transport stream. Again, since a transport stream is a serial stream of multiplexed audio and video packets, the "parallel" nature of the audio/video elementary streams (substreams) in each transport stream is only illustrative of the audio/video lip sync of each transport stream.

FIG. 2 illustrates the splicing operation of a first transport stream 210 with a second transport stream 220 to produce an output stream 230.



More specifically, the first three (3) video access units  $212_{1,3}$  and their associated four (4) audio access units  $214_{1,4}$ , are initially spliced from the first transport stream 210 to the output stream 230. Next, the first three (3) video access units  $222_{1,3}$  and their associated four (4) audio access units  
5  $224_{1,4}$ , are spliced from the second transport stream 220 to the output transport stream 230 behind those access units taken from the first transport stream 210. In FIG. 2, the splicing operation is premised on video frame alignment (picture presentation continuity), i.e., video access unit  $222_1$  follows immediately after video access unit  $212_3$ .

10 To preserve the audio/video lip sync in the splicing operation, it is important to carefully analyze the timing relationship in the transport streams. An MPEG compliant transport stream comprises a plurality of packets with each packet having a length of 188 bytes, 4 bytes of header and 184 bytes of payload. The program clock reference (PCR) fields are carried  
15 in the adaptation field that follows the 4-byte transport packet header. The value encoded in the PCR field indicates the time  $t(i)$ , where  $i$  refers to the byte containing the last bit of the PCR base field. The PCR values are derived from the 27 MHz reference clock 150 and are inserted into the transport stream to effect synchronization between the encoding and  
20 decoding process. More importantly, the PCR also serves as the basis for the generation of the presentation time stamp (PTS) and the decoding time stamp (DTS), that represent the presentation and decoding time of audio and video signals, respectively. These time stamps play an important role in audio/video lip sync.

25 The relationship between PCR and the PTS/DTS is generally defined for each transport stream at the time of creation of the stream. However, these timing information contain differences when compared to other PCR and PTS/DTS in other transport streams stored within server 114, or when compared to the timing of "real" time encoding of the encoding  
30 section 120 as illustrated in the digital studio. This timing discrepancy between multiple input transport streams must be addressed by the digital studio as the studio attempts to splice different transport streams into a single output transport stream. An example of a solution to address such timing discrepancy is disclosed in an accompanying patent application  
35 filed simultaneously herewith on \_\_\_\_\_ with the title "Timing

Correction Method and Apparatus" (attorney docket SAR 12389; serial number \_\_\_\_\_), hereby incorporated by reference.

Returning to FIG. 2, each of the input transport streams 210 and 220 contains regularly spaced PTSs for both the audio and video access units. However, the PTSs for the video access units may or may not be time aligned with the PTSs from those of the associated audio access units. Furthermore, the video access units for the second transport stream 220, after the splice point 250, also contain regularly spaced PTSs but they are offset from those video access units of the first transport stream 210 by any amount within the range of possible values for PTSs. To maintain picture presentation continuity when the splicing operation is completed, the PTS for the first video access unit after the splice can be re-stamped with a value that would be expected after the end of the last video access unit before the splice operation. This process, as disclosed in accompanying application SAR 12389, maintains picture presentation continuity as long as a decoder buffer (not shown) is properly managed as recommended in the MPEG standards.

In brief, the presentation time for the video access unit following the splice point 250 can be calculated using the frame rate and the last PTS before the splice. A relationship between the calculated PTS and the original PTS stamped in the stream following the splice exists in the form of an offset. This offset can be applied to each PTS in the stream after the splice. The PTS for the audio access unit after the splice can be found by using the same offset that was used in the video stream. This process maintains the original synchronization between the video and the audio elementary streams both before and after the splice.

Returning to FIG. 2, in order to avoid the overlapping of audio access units 214, and 224, in the output transport stream 230, the audio access units 214, is deleted from the output transport stream 230 during the splicing operation. The detection of the pending overlapping condition is performed by the lip sync evaluator 130. The method of detection can be summarized as follows:

$$\text{if } A1 + B \leq A2 \pm X, \text{ then no overlap condition} \quad (1)$$

if  $A1 + B > A2 \pm X$ , then overlap condition exists (2)

where A1 represents a time stamp, e.g., PTS, for the last audio access unit  
5 (e.g., audio access unit 214<sub>4</sub>) in the first transport stream. A2 represents a  
time stamp, e.g., PTS, for the first audio access unit (e.g., audio access  
unit 224<sub>1</sub>) in the second transport stream. B represents the duration of the  
last audio access unit, e.g., 32 msec. in accordance with the ATSC A/53  
standard. X represents a general timing offset between the first transport  
10 stream 210 and the second transport stream 220, where X can be  
determined as disclosed above in the accompanying application, with  
attorney docket SAR 12389.

The offset X can be calculated by an audio/video PTS/DTS retiming  
section 137 as illustrated in FIG. 1, where the calculated offset X is passed  
15 to the lip sync evaluator 130 for use as discussed above. However, if the  
two input transport streams are already time synchronized, e.g.,  
previously adjusted within the server 114 or the two transport streams  
were originally created with the same timebase, then the offset X should  
be zero.

20 To illustrate, FIG. 2 depicts the deletion of the audio access unit 214<sub>4</sub>  
from the output transport stream 230. As shown pictorially, audio access  
unit 214<sub>4</sub> will overlap with audio access unit 224<sub>1</sub> in the output transport  
stream. More specifically, the PTS of the audio access unit 214<sub>4</sub> is  
illustrated as having a value that is trailing the PTS of the video access  
25 unit 212<sub>3</sub>. Namely, the video signal contained within the video access unit  
212<sub>3</sub> will be displayed ahead in time before the presentation of the audio  
signal contained within the audio access unit 214<sub>4</sub>. This creates the effect  
where the duration of the audio signal contained within the audio access  
unit 214<sub>4</sub> overlaps with the start of presentation of the audio signal  
30 contained within the audio access unit 224<sub>1</sub>. If this condition is detected,  
the present invention deletes the audio access unit 214<sub>4</sub> prior to the  
splicing operation, thereby leaving a gap 240 in the audio elementary  
stream. Such gaps in audio signal are typically handled by the audio  
decoder (not shown) by gracefully muting the audio output or adjusting  
35 the volume of the audio output.

Although the preferred embodiment deletes the last audio access unit 214<sub>4</sub>, it should be understood that the overlapping condition can be resolved by deleting the first audio access unit 224<sub>1</sub> from the second transport stream 220. The judgment as to which audio access unit to delete is an artistic decision. In certain applications, it may be desirable to always present the newly spliced video access units with their associated audio access units, while other applications may prefer the presentation of the audio signal from the previous transport stream to be completed.

FIG. 3 illustrates a splicing operation where there is no overlap of audio access units. More specifically, FIG. 3 illustrates a splicing operation of a first transport stream 310 with a second transport stream 320 to produce an output stream 330. The first three (3) video access units 312<sub>1,3</sub> and their associated four (4) audio access units 314<sub>1,4</sub>, are initially spliced from the first transport stream 310 to the output stream 330. Next, the first three (3) video access units 322<sub>1,3</sub> and their associated four (4) audio access units 324<sub>1,4</sub>, are spliced from the second transport stream 320 to the output transport stream 330, behind those access units taken from the first transport stream 310.

Similarly, the splicing operation in FIG. 3 is also premised on maintaining picture presentation continuity, i.e., video access unit 322<sub>1</sub> follows immediately after video access unit 312<sub>3</sub>. However, unlike the splicing operation illustrated in FIG. 2, audio access unit 314<sub>4</sub> does not overlap audio access unit 324<sub>1</sub>. Absence an overlap condition, the second transport stream is simply spliced into the output transport stream without the need to delete any audio access unit.

FIG. 4 illustrates an alternative embodiment of the present invention where the splicing operation is premised on maintaining audio presentation continuity. More specifically, FIG. 4 illustrates a splicing operation of a first transport stream 410 with a second transport stream 420 to produce an output stream 430. The first three (3) video access units 412<sub>1,3</sub> and their associated four (4) audio access units 414<sub>1,4</sub>, are initially spliced from the first transport stream 410 to the output stream 430. Next, the first three (3) video access units 422<sub>1,3</sub> and their associated four (4) audio access units 424<sub>1,4</sub>, are spliced from the second transport stream 420 to the output transport stream 430, behind those access units taken from

the first transport stream 410.

However, unlike FIGs. 2-3, the splicing operation is premised on maintaining audio presentation continuity, i.e., audio access unit 424<sub>1</sub> follows immediately after audio access unit 414<sub>4</sub>. This creates a potential  
5 overlap of video access units, e.g., video access unit 412<sub>3</sub> will overlap with video access unit 422<sub>1</sub> in the output transport stream. If this condition is detected, the present invention deletes the video access unit 412<sub>3</sub> prior to the splicing operation, thereby leaving a gap 440 in the video elementary  
10 stream. Such gaps in the video signal maybe handled by future video decoder (not shown) by gracefully changing the frame or field rate at the discontinuity. Again, artistic decision may decide which video access unit to delete.

Furthermore, the above equations (1) and (2) can be similarly applied where the splicing operation is premised on maintaining audio  
15 presentation continuity. More specifically, A1 represents a time stamp for said last video access unit of the first transport stream. A2 represents a time stamp for said first video access unit of the second transport stream. B represents a duration of said last video access unit of the first transport stream, and X represents a timing offset between said first transport  
20 stream and said second transport stream.

FIG. 5 illustrates a flowchart of a method 500 for preserving audio/video lip sync when splicing a plurality of data streams into an output stream. Referring to FIG. 5, the method 500 starts at step 505 and proceeds to step 510, where method 500 determines whether the current  
25 splicing operation is the start of a new output stream or whether a data stream is spliced into an existing output stream. If the query is negatively answered, method 500 proceeds to step 520. If the query at step 510 is affirmatively answered, method 500 proceeds to step 515 where a first data stream (or portion thereof) is selected (spliced) to be the start of a new  
30 output stream.

In step 515, a first data stream, e.g., a first transport stream (or portion thereof), is spliced to form the beginning of an output stream. In the preferred embodiment, the selected data stream passes through (temporally stored within) a buffer, e.g., within memory 132, before being  
35 forwarded to the service multiplexer 135. Generally, the buffer is

implemented as a First-In-First Out (FIFO) buffer. This buffering of the selected data streams creates a time window, where the lip sync evaluator 130 is allowed to detect for overlapping of access units as described below.

It should be understood that the selected data streams can be stored  
5 in a buffer for analysis prior to being spliced into the output stream for certain applications. However, storing such data streams requires large buffer size and may not be well suited for real time applications. Furthermore, since an overlapping condition generally occurs at the splice point, method 500 generally only needs to analyze the various access  
10 units that are proximate to the splice point.

In step 520, a second data stream, e.g., a second transport stream (or portion thereof), is selected and spliced at a splice point behind the first selected data stream to continue the formation of the output stream. Similarly, the second selected data stream also passes through a buffer,  
15 e.g., within memory 132, before being forwarded to the service multiplexer 135 as illustrated in FIG. 1.

In step 525, method 500 determines whether an overlapping condition exists at the splice point with regard to audio or video access units. If the query is negatively answered, method 500 proceeds to optional  
20 step 535. If the query at step 510 is affirmatively answered, method 500 proceeds to step 530 where one of the overlapping access unit is deleted to resolve the overlapping condition.

The deletion step 530 can be implemented within the buffer 132 or the selected overlapping access unit can be simply dropped without being  
25 forwarded to the service multiplexer 135. Furthermore, the decision whether to delete an audio or a video access unit is premised on the alignment scheme selected for an application as discussed above.

In optional step 535, if the selected data streams are stored and analyzed in their entirety before the actual splice operation, then the  
30 selected data streams are spliced together in this step. Otherwise, this step can be omitted if the data streams are spliced and analyzed "on the fly" as discussed above.

In step 540, method 500 determines whether additional data streams are scheduled to be spliced into the output stream. If the query at  
35 step 540 is affirmatively answered, method 500 proceeds to step 520, where

the next data stream is selected and spliced into the output stream. If the query is negatively answered, method 500 ends in step 545.

There has thus been shown and described a novel method and apparatus for preserving audio/video lip sync when splicing data streams.

- 5 Many changes, modifications, variations and other uses and applications of the subject invention will, however, become apparent to those skilled in the art after considering this specification and the accompanying drawings which disclose the embodiments thereof. All such changes, modifications, variations and other uses and applications which do not  
10 depart from the spirit and scope of the invention are deemed to be covered by the invention, which is to be limited only by the claims which follow.

What is claimed is:

1. A method for preserving lip sync during a splicing operation where a plurality of data streams is multiplexed into an output stream, where  
5 each of the data streams contains an audio stream having a plurality of audio access units and a video stream having a plurality of video access units that are in lip sync, said method comprising the steps of:
- (a) splicing a portion of a first data stream into the output stream;
  - (b) splicing a portion of a second data stream into the output stream;
  - 10 (c) determining at a splice point whether an overlap condition exists between a last audio access unit from said first data stream and a first audio access unit from said second data stream; and
  - (d) deleting either said last audio access unit or said first audio access unit if an overlap condition exists.
- 15
2. The method of claim 1, wherein said data streams are transport streams.
3. The method of claim 1, wherein said determining step (c)  
20 determines said overlapping condition in accordance with:

if  $A1 + B \leq A2 \pm X$ , then no overlap condition

if  $A1 + B > A2 \pm X$ , then overlap condition exists

25

where A1 represents a time stamp for said last audio access unit, A2 represents a time stamp for said first audio access unit, B represents a duration of said last audio access unit, and X represents a timing offset between said first data stream and said second data stream.

30

4. A method for preserving lip sync during a splicing operation where a plurality of data streams is multiplexed into an output stream, where each of the data streams contains an audio stream having a plurality of audio access units and a video stream having a plurality of video access  
35 units that are in lip sync, said method comprising the steps of:



- (a) splicing a portion of a first data stream into the output stream;
- (b) splicing a portion of a second data stream into the output stream;
- (c) determining at a splice point whether an overlap condition exists between a last video access unit from said first data stream and a first  
5 video access unit from said second data stream; and
- (d) deleting either said last video access unit or said first video access unit if an overlap condition exists.

5. The method of claim 4, wherein said data streams are transport  
10 streams.

6. The method of claim 4, wherein said determining step (c) determines said overlapping condition in accordance with:

15 if  $A1 + B \leq A2 \pm X$ , then no overlap condition

if  $A1 + B > A2 \pm X$ , then overlap condition exists

where  $A1$  represents a time stamp for said last video access unit,  $A2$   
20 represents a time stamp for said first video access unit,  $B$  represents a duration of said last video access unit, and  $X$  represents a timing offset between said first data stream and said second data stream.

7. An apparatus (100) for preserving lip sync during a splicing  
25 operation where a plurality of data streams (210, 220, 310, 320, 410, 420) is multiplexed into an output stream (230, 330, 430), where each of the data streams contains a first substream (214, 224, 314, 324, 414, 424) having a plurality of access units and a second substream (212, 222, 312, 322, 412, 422) having a plurality of access units, said apparatus comprising:

30 a splicer (135) for splicing a portion of a first data stream and a portion of a second data stream into the output stream;

a means (130), coupled to said splicer, for determining at a splice point whether an overlap condition exists between said access units from said first data stream and said second data stream; and

-16-

a means (130), coupled to said determining means, for deleting one or more of said access units if an overlap condition exists.

8. The apparatus of claim 7, wherein said first substream (214, 224,  
5 314, 324, 414, 424) is an audio stream having a plurality of audio access units and said second substream (212, 222, 312, 322, 412, 422) is a video stream having a plurality of video access units that are in lip sync to said audio access units.

10 9. The apparatus of claim 8, wherein said deleted access unit is an overlapping audio access unit.

10. The apparatus of claim 8, wherein said deleted access unit is an overlapping video access unit.

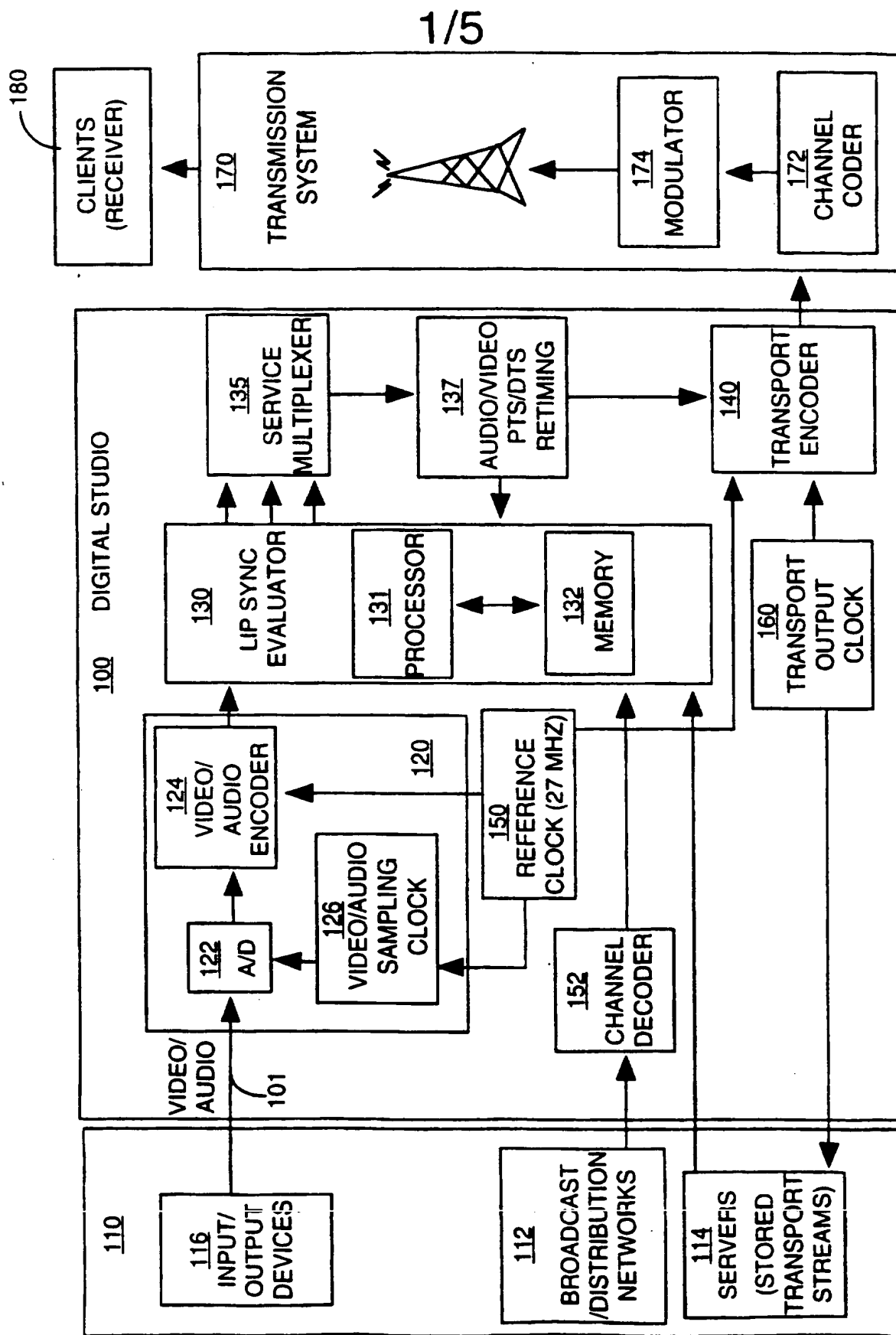


FIG. 1

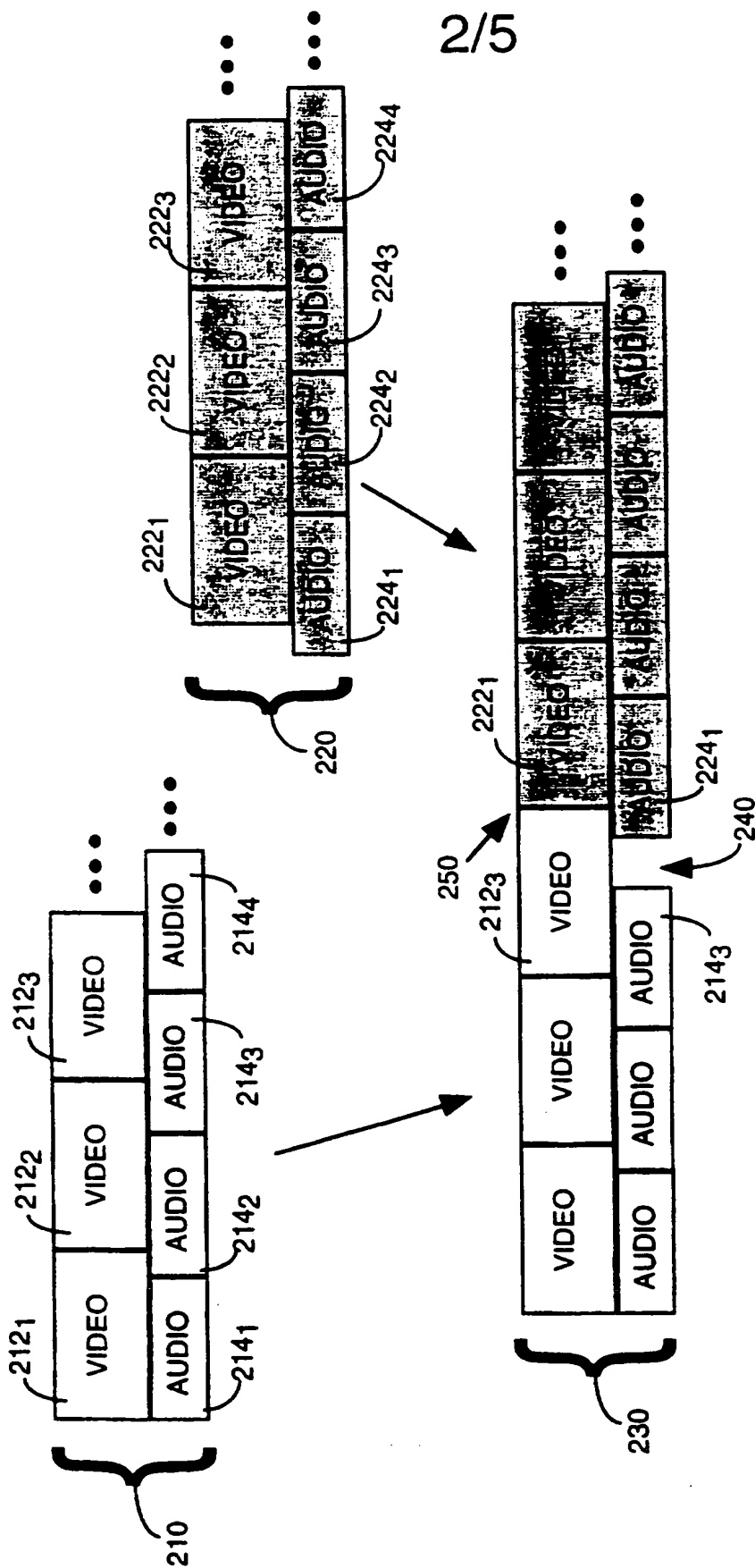


FIG. 2

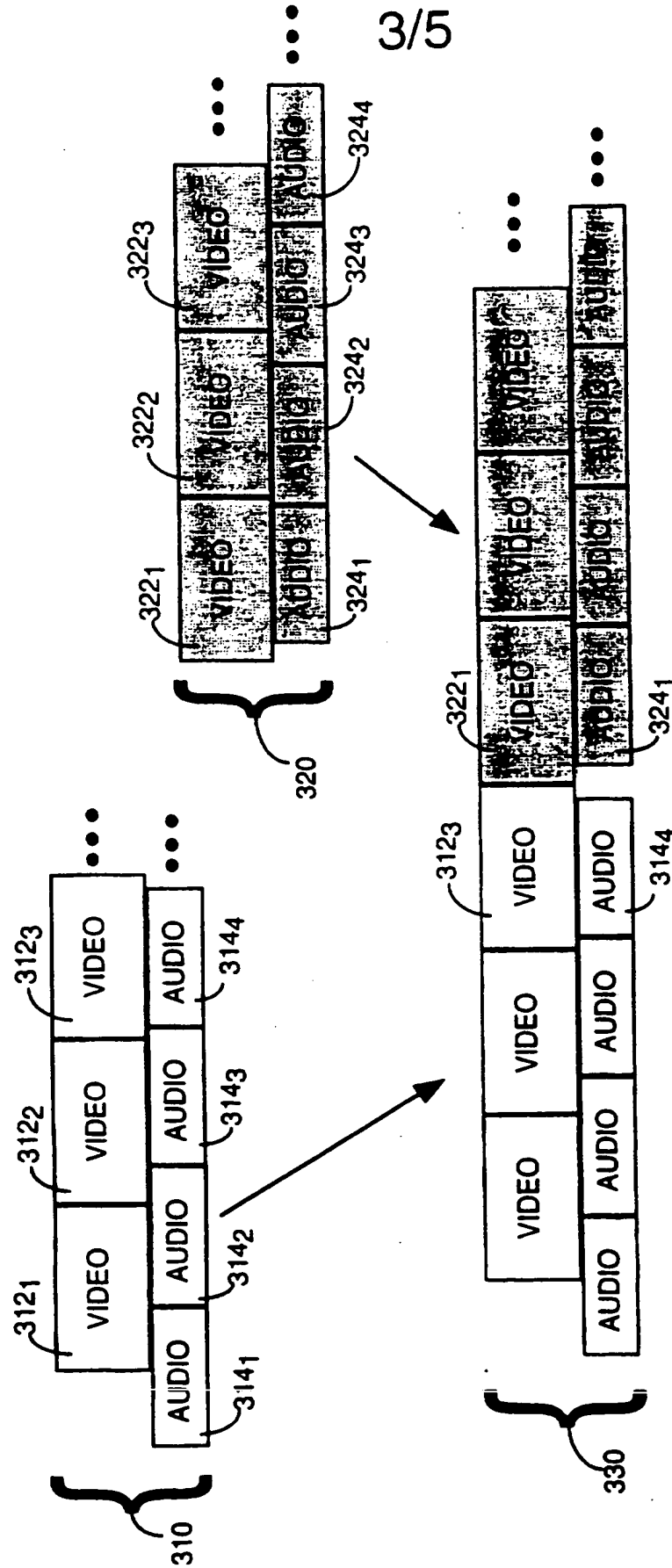


FIG. 3

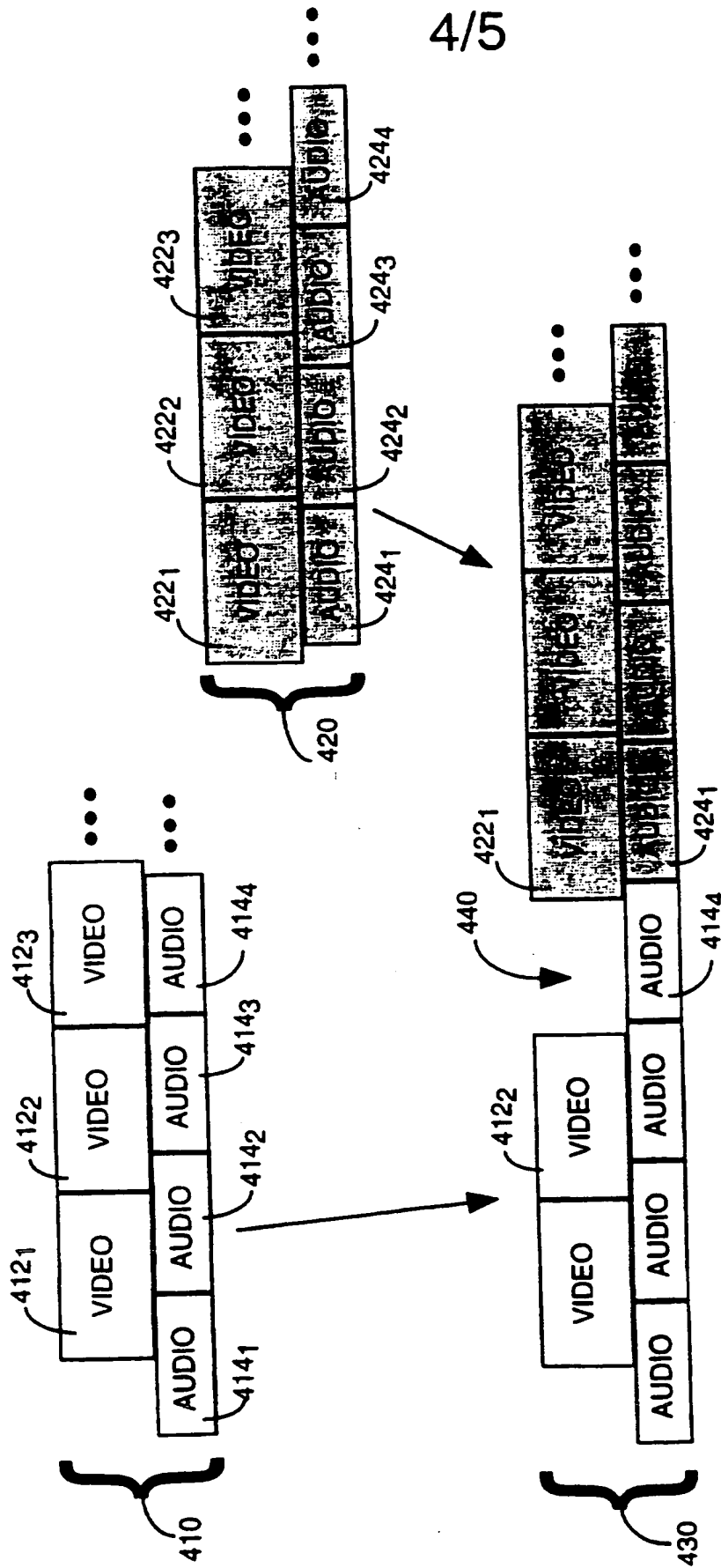
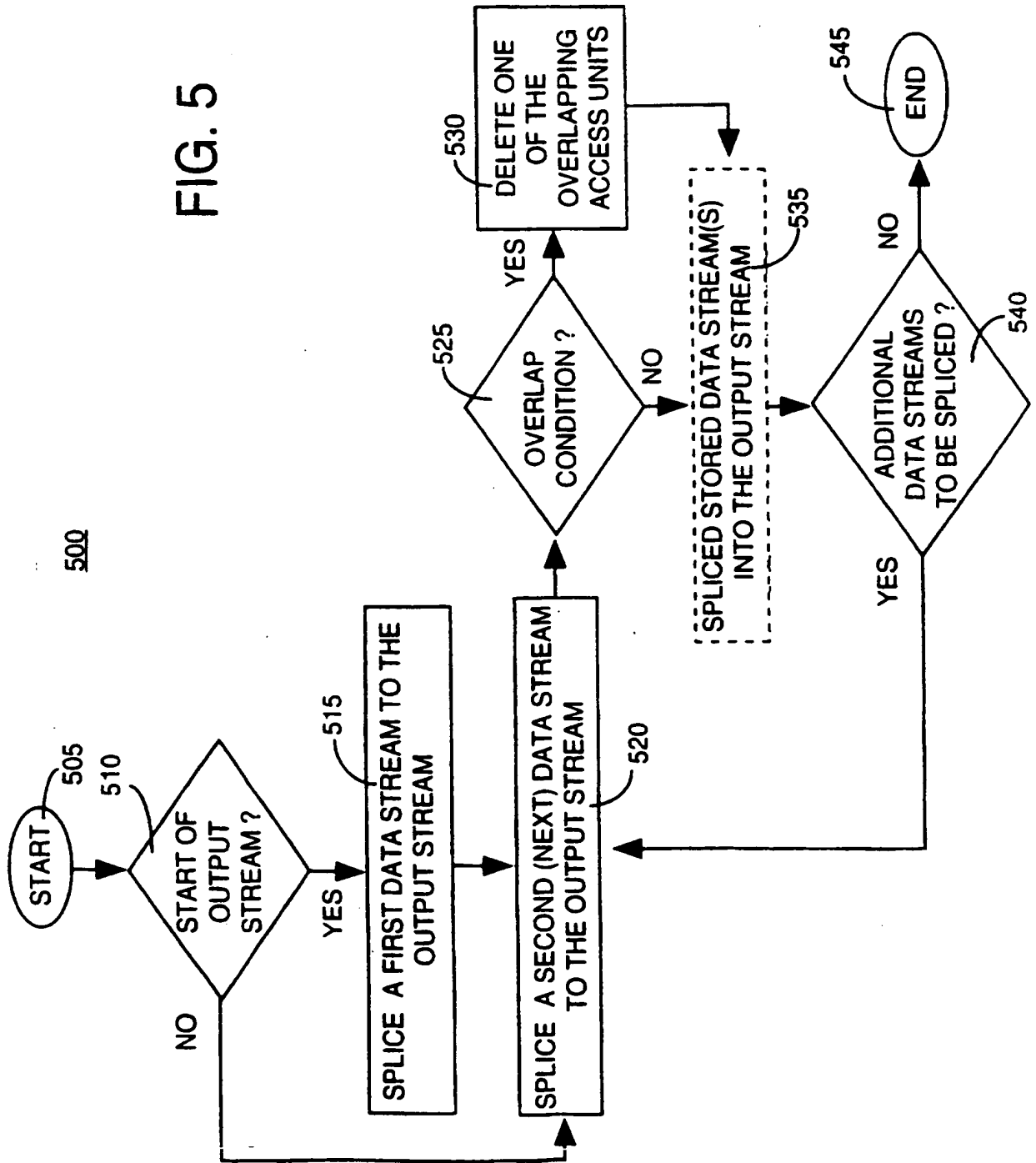


FIG. 4

5/5

FIG. 5



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/10186

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) H04N 9/475

US CL 348/512, 515, 518

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. 348/512, 515, 518, 517, 516; 360/ 19.1, 36.1, 9.1; 352/12, 25

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,387,943 A (SILVER) 07 FEBRUARY 1995, COL. 2, LINE 39 TO COL. 4, LINE 8.	1-10
A	US, 5,430,485 A (LANKFORD ET AL.) 04 JULY 1995, COL. 6, LINES 47-68.	1-10
A	US 4,703,355 A (COOPER) 27 OCTOBER 1987, COL. 2, LINE 58 TO COL. 3, LINE 42.	1-10

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	*A* document member of the same patent family

Date of the actual completion of the international search

23 AUGUST 1997

Date of mailing of the international search report

30 OCT 1997

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

GLEN BURGESS

Telephone No. (703) 305-3900



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

### **IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**